IN THE CLAIMS:

Please cancel claims 4 and 18.

1   1.   (Resubmitted)   A method for clustering a set S of n data points to find k

2   final centers, comprising:

3   partitioning said set S into P disjoint pieces $S_1,...,S_P$;

4   for each said piece S_i, determining a set D_i of k intermediate centers;

5   assigning each data point in each piece S_i to the nearest one of said k intermediate

6   centers;

7   weighting each of said k intermediate centers in each set D_i by the number of

8   points in the corresponding piece S_i assigned to that center; and

9   clustering said weighted intermediate centers together to find said k final centers,

10   said clustering performed using a specific error metric and a clustering method A.

1   2.   (Currently Amended) ~~A~~ The A method according to claim 1 further

2   comprising:

3   merging said weighted centers into a single dataset D' prior to clustering.

1   3.   (Currently Amended) ~~A~~ The method according to claim 1 wherein P is

2   sufficiently large enough such that each piece S_i obeys the constraint $|S_i| < M$, where M is

3   the size of a physical memory or a portion thereof to be used in processing said each

4   piece.

1   4.   (Canceled)   A method according to claim 1 wherein if P is not

2   sufficiently large enough such that each piece S_i obeys the constraint $|S_i| < M$, where M is

3   the size of a physical memory or a portion thereof to be used in processing said each

4   piece, then iteratively performing partitioning, determining, assigning, and weighting until

3

5   the sets D' of weighted intermediate centers generated thereby obeys the constraint |D'| <

6   M.

1   5.   (Currently Amended) ~~A~~ The method according to claim 4 wherein said

2   clustering is performed upon iteratively obtained weighted intermediate clusters.

1   6.   (Currently Amended) ~~A~~ The method according to claim 4 wherein said set

2   S is replaced by weighted intermediate centers of the previous iteration when iteratively

3   performing said partitioning, determining, assigning, and weighting.

1   7.   (Currently Amended) ~~A~~ The method according to claim 1 wherein said

2   determining is performed using said specific error metric and said clustering method A.

1   8.   (Currently Amended) ~~A~~ The method according to claim 1 wherein said

2   specific error metric is the minimizing of the sum of the squares of the distances between

3   points and their nearest centers.

1   9.   (Currently Amended) ~~A~~ The method according to claim 1 wherein said

2   specific error metric is the minimizing of the sum of the distances between points and

3   their nearest centers.

1   10.   (Currently Amended) ~~A~~ The method according to claim 1 wherein said

2   clustering method is an approximation-based method.

1   11.   (Currently Amended) ~~A~~ The method according to claim 8 wherein the

2   distance is the Euclidean distance.

1   12.   (Currently Amended) ~~A~~ The method according to claim 9 wherein the

2   distance is the Euclidean distance.

1   13.   (Currently Amended) ~~A~~ The method according to claim 1 further

2   comprising:

3   considering a second set of data points for obtaining a second k final centers after

4   said set S is clustered;

5   repeating partitioning, determining, assigning and weighting for said second set of

6   data points; and

7   clustering weighted intermediate centers obtained from said second set of data

8   points together with said weighted intermediate centers obtained from said data set S, said

9   clustering performed using said specific error metric and said clustering method A.

1   14.   (Currently Amended) ~~A~~ The method according to claim 1 wherein said

2   partitioning, determining, assigning and weighting is performed in parallel for each piece

3   S_i.

1   15.   (Resubmitted)   An article comprising a computer readable medium

2   having instructions stored thereon which when executed causes clustering a set S of n data

3   points to find k final centers, said clustering implemented by:

4   partitioning said set S into P disjoint pieces $S_1,...,S_P$;

5   for each said piece S_i, determining a set D_i of k intermediate centers;

6   assigning each data point in each piece S_i to the nearest one of said k intermediate

7   centers;

8   weighting each of said k intermediate centers in each set D_i by the number of

9   points in the corresponding piece S_i assigned to that center; and

5

10   clustering said weighted intermediate centers together to find said k final centers,

11   said clustering performed using a specific error metric and a clustering method A.

1   16.   (Currently Amended) ~~An~~ The article according to claim 15 further

2   implemented by:

3   merging said weighted centers into a single dataset D' prior to clustering.

1   17.   (Currently Amended) ~~An~~ The article according to claim 15 wherein P is

2   sufficiently large enough such that each piece S_i obeys the constraint $|S_i| < M$, where M is

3   the size of a physical memory or a portion thereof to be used in processing said each

4   piece.

1   18.   (Canceled)   The article according to claim 15 wherein if P is not

2   sufficiently large enough such that each piece S_i obeys the constraint $|S_i| < M$, where M is

3   the size of a physical memory or a portion thereof to be used in processing said each

4   piece, then iteratively performing partitioning, determining, assigning, and weighting until

5   the sets D' of weighted intermediate centers generated thereby obeys the constraint $|D'| <$

6   M.

1   19.   (Currently Amended) ~~An~~ The article according to claim 1 further

2   implemented by:

3   considering a second set of data points for obtaining a second k final centers after

4   said set S is clustered;

5   repeating partitioning, determining, assigning and weighting for said second set of

6   data points; and

6

7   clustering weighted intermediate centers obtained from said second set of data

8   points together with said weighted intermediate centers obtained from said data set S, said

9   clustering performed using said specific error metric and said clustering method A,

10   resulting in said second k final clusters.

1   20.   (Currently Amended) ~~An~~ The method according to claim 1 wherein said

2   partitioning, determining, assigning and weighting is performed in parallel for each piece

3   S_i.

1   21.   (Resubmitted) An apparatus for clustering a set S of n data points to find k

2   final centers, said apparatus comprising:

3   a main memory;

4   a processor coupled to said memory, said processor configured to partition said set

5   S into P disjoint pieces $S_1,...,S_P$ such that each piece S_i fits in main memory, said each

6   piece S_i first stored separately in said main memory and then clustered by said processor

7   performing:

8   for each said piece S_i, determining a set D_i of k intermediate centers;

9   assigning each data point in each piece S_i to the nearest one of said k intermediate

10   centers;

11   weighting each of said k intermediate centers in each set D_i by the number of

12   points in the corresponding piece S_i assigned to that center; and

13   clustering said weighted intermediate centers together to find said k final centers,

14   said clustering performed using a specific error metric and a clustering method A.

22.   (Resubmitted) An apparatus for clustering a set S of n data points to find k final centers, said apparatus comprising:

a main memory;

a plurality of processors coupled to said main memory, one of said processors configured to partition said set S into P disjoint pieces $S_1,...,S_P$ such that each piece S_i fits in main memory, said each piece S_i first stored separately in said main memory and then clustered by each said processor performing:

for each said piece S_i, determining a set D_i of k intermediate centers;

assigning each data point in each piece S_i to the nearest one of said k intermediate centers; and

weighting each of said k intermediate centers in each set D_i by the number of points in the corresponding piece S_i assigned to that center, further wherein after aid weighting, one of said processors finally clustering said weighted intermediate centers together to find said k final centers, said clustering performed using a specific error metric and a clustering method A.